

在开发多媒体作品时,数字音频信息是经常采用的元素。数字音频信息主要的表现形式是讲解、声效和音乐。通过这些媒介,能烘托多媒体作品的主题并营造气氛,尤其对于多媒体教学光盘、多媒体广告、多媒体课件等,数字音频信息显得更加重要。

本章主要内容:

- 声音的基本知识;
- 声音信号数字化技术;
- 数字音频文件格式;
- 语音识别技术;
- 基于 Adobe Audition 的数字音频处理技术。

3.1 基本知识

声音是人们用来传递信息、交流感情最方便、最熟悉的方式之一。本节介绍声音的概念以及声音的三要素。

3.1.1 认识声音

声音是通过物体振动产生的。声音是通过介质(空气或固体、液体)传播并能被人或动物听觉器官所感知的波动现象。

声音是一种压力波。当演奏乐器、拍打一扇门或者敲击桌面时,它们的振动会引起介质——空气分子——有节奏的振动,使周围的空气产生疏密变化,形成疏密相间的纵波,这就产生了声波,这种现象会一直延续到振动消失为止。

声音的三个重要指标如下。

- (1) 振幅: 声波的高低幅度,表示声音的强弱。
- (2) 周期: 两个相邻声波之间的时间长度。
- (3) 频率: 频率是每秒振动的次数,它的单位为赫(Hz)。

3.1.2 声音的三要素

从听觉角度看,声音具有音调、音色和音强三要素。声音质量的高低主要取决于这三要素。

1. 音调

声音的高低称为音调。音调与声音的频率有关,声源振动的频率越



高,声音的音调就越高;声源振动的频率越低,声音的音调就越低。通常把音调高的声音叫高音,音调低的声音叫低音。

2. 音色

音色指声音的感觉特性,与声音波形相关。声音分为纯音和复音两种类型。所谓纯音,是指振幅和周期均为常数的声音;所谓复音,是指具有不同频率和不同振幅的混合声音。大自然中的声音大部分是复音。在复音中,最低频率的声音是“基音”,它是声音的基调。其他频率的声音称为“谐音”,也叫泛音。基音和谐音是构成声音音色的重要元素。若中、高频谐音丰富,音色就明亮,如小号;若低频谐音丰富,音色就低沉,如低音贝司。

3. 音强

音强就是声音的强度,也被称为声音的响度,常说的“音量”也是指音强。音强与声波的振幅成正比,振幅越大,音强越大。如果要改变原始声音的音强,在把声音数字化以后,可以使用音频处理软件进行编辑处理。

3.2 声音信号数字化技术

声音是自然界中一切可以听到的振动波,为了用计算机表示和处理声音,必须把声音进行数字化。数字化了的声音叫做“数字音频信号”。

3.2.1 声音信号数字化过程

把模拟信号转换成数字信号的过程称为模/数转换,它主要包括采样、量化和编码三个步骤,如图 3-1 所示。

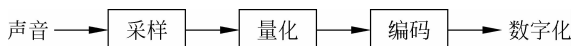


图 3-1 声音信号数字化过程

采样: 在时间轴上对信号数字化,即每隔相等的一段时间在声音信号波形曲线上采集一个信号样本。

量化: 在幅度轴上对信号数字化。采样后的数值不一定就能在计算机内部被表示,因此必须将每一个样本值归入预先编排的最近的量化级上。如果幅度划分是等间隔的,就称为线性量化,否则就称为非线性量化。

编码: 按一定格式记录采样和量化后的数字数据。

脉冲编码调制(Pulse Code Modulation, PCM)是一种模数转换的最基本编码方法,CD-DA 就是采用的这种编码方式。

3.2.2 数字音频的技术指标

数字化后的声音质量各有不同,声音质量的好坏主要取决于数字化过程中的采样频率、量化位数、声道数和编码算法等技术指标。

1. 采样频率

采样频率是指一秒钟内采样的次数。采样频率的选中应该遵循奈奎斯特(Harry Nyquist)采样理论:如果对某一模拟信号进行采样,则采样后可还原的最高信号频率只有

采样频率的一半,或者说只要采样频率高于输入信号最高频率的两倍,就能从采样信号系列重构原始信号。

根据该采样理论,CD激光唱盘采样频率为44kHz,可记录的最高音频为22kHz,这样的音质与原始声音相差无几,也就是常说的超级高保真音质(super High Fidelity, HiFi)。采样的三个标准频率分别为44.1kHz,22.05kHz和11.025kHz。

2. 量化位数

量化位数是对模拟音频信号的幅度进行数字化,它决定了模拟信号数字化以后的动态范围。由于计算机按字节运算,一般的量化位数为8位和16位。量化位越高,信号的动态范围越大,数字化后的音频信号就越可能接近原始信号,但所需要的存储空间也越大。

3. 声道数

声道数指的是一次同时产生的声波组数。数字化音频一般有单声道和双声道之分。双声道又称为立体声,在硬件中要占两条线路,音质、音色好,但立体声数字化后所占存储空间比单声道多一倍。

专家点拨:除了单声道和双声道以外,还有环绕4声道、5.1声道、7.1声道等。

4. 编码算法

编码的作用主要有两个,一是采用一定的格式来记录数字数据,二是采用一定的算法来压缩数字数据。压缩编码的基本指标之一就是压缩比,它通常小于1。音频数据压缩比的表达式为:

$$\text{音频数据压缩比} = \frac{\text{压缩后的音频数据}}{\text{压缩前的音频数据}}$$

压缩算法包括有损压缩和无损压缩。有损压缩指解压后数据不能完全复原,要丢失一部分信息。压缩比越小,丢掉的信息越多、信号还原后失真越大。根据不同的应用,可以选用不同的压缩编码算法,如PCM、ADPC、MP3、RA等。

3.2.3 数字音频的音质与数据量

这里讨论的数字音频是指WAV格式的波形音频文件。数字音频的声音质量好坏,取决于采样频率的高低、量化位数和声道数。如表3-1所示,列出了声音质量与各种数字音频技术指标的对应关系。从表中可以看出,音质越好,音频文件的数据量越大。

表3-1 音质与数字音频技术指标的关系

采样频率/kHz	量化位数	声道数	数据量/(KB/s)	音频质量
8	8	单声道	8	一般质量
	8	双声道	16	
	16	单声道	16	
	16	双声道	31	
11.025	8	单声道	11	电话质量
	8	双声道	22	
	16	单声道	22	
	16	双声道	43	



续表

采样频率/kHz	量化位数	声道数	数据量/(KB/s)	音频质量
22.05	8	单声道	22	收音质量
	8	双声道	43	
	16	单声道	43	
	16	双声道	86	
44.1	8	单声道	43	
	8	双声道	86	
	16	单声道	86	
	16	双声道	172	CD质量

表 3-1 中的数据量由下列算式计算出:

$$\text{数据量} = (\text{采样频率} \times \text{量化位数} \times \text{声道数}) / 8$$

例如,CD 质量的数字音频每秒钟的数据量为:

$$\text{数据量} = (44\ 100\text{Hz} \times 16\text{b} \times 2) \div 8 = 176\ 400\text{B} (\text{约为 } 172\text{KB})$$

如果以 CD 质量(44.1kHz 的采样频率、16 位的量化位数、双声道立体声)记录一首时长 5min(300s)的音乐,那么数据量为:

$$172\text{KB/s} \times 300\text{s} = 51\ 600\text{KB} (\text{约为 } 50.39\text{MB})$$

由计算结果可以看出,音频文件的数据量问题十分重要。为了节省存储空间,通常在保证基本声音质量的前提下,可以采用稍低一些的采样频率。一般情况下,人的语音采用 11.025kHz 的采样频率、8b、单声道已经足够;如果是音乐,22.05kHz 的采样频率、8b、双声道立体声基本可以满足需要。

3.3 数字音频文件格式

目前,随着多媒体技术的发展,数字音频格式也层出不穷,下面介绍几种常见的数字音频格式以及这些格式的相互转换方法。

3.3.1 各种音频格式介绍

1. CD 格式

CD 是当今世界上音质最好的音频格式。在大多数播放软件的“打开文件类型”对话框中,都可以看到 cda 格式,这就是 CD 音轨。标准 CD 格式是 44.1kHz 的采样频率,速率 88KB/s,16 位量化位数。

因为 CD 音轨可以说是近似无损的,所以它的声音基本上是忠于原声的。如果是音响发烧友,CD 是首选,它能让人感受到天籁之音。CD 光盘可以在 CD 唱机中播放,也能用计算机里的各种播放软件来播放。

一个 CD 音频文件是一个 cda 文件,但这只是一个索引信息,并不是真正包含的声音信息,因此不论 CD 音乐的长短,在计算机上看到的 cda 文件都是 44B 长。

专家点拨: 不能直接复制 CD 格式的 cda 文件到硬盘上播放,需要使用像 EAC 这样的抓音轨软件把 CD 格式的文件转换成 WAV 格式的文件,这个转换过程如果光盘驱动器质

量过关而且 EAC 的参数设置得当,可以说基本上是无损抓音频。

2. WAV 格式

WAV 是微软公司开发的一种声音文件格式,它符合 PIFResource Interchange File Format 文件规范,用于保存 Windows 平台的音频信息资源,被 Windows 平台及其应用程序所支持。这种音频格式文件的扩展名为 WAV。

WAV 格式支持 MSADPCM、CCITT A LAW 等多种压缩算法,支持多种音频位数、采样频率和声道,标准格式的 WAV 文件和 CD 格式一样,也是 44.1kHz 的采样频率,速率 88K/s,16 位量化位数。WAV 格式的声音文件质量和 CD 相差无几,也是目前 PC 上广为流行的声音文件格式,几乎所有的音频编辑软件都“认识”WAV 格式。

专家点拨: 由苹果公司开发的 AIFF(Audio Interchange File Format)和为 UNIX 系统开发的 AU 格式,它们都和 WAV 非常相像,在大多数的音频编辑软件中也都支持这几种常见的音频格式。

3. MP3 格式

MP3 格式诞生于 20 世纪 80 年代的德国,所谓的 MP3 也就是指 MPEG 标准中的音频部分,即 MPEG 音频层。根据压缩质量和编码处理的不同分为三层,分别对应 MP1、MP2、MP3 这三种声音文件。

需要提醒读者注意的是,MPEG 音频文件的压缩是一种有损压缩,MPEG3 音频编码具有(10:1)~(12:1)的高压缩率,同时基本保持低音频部分不失真,但是牺牲了声音文件中 12~16kHz 高频部分的质量来换取文件的尺寸,相同长度的音乐文件,用 MP3 格式来储存,一般只有 WAV 文件的 1/10,而音质要次于 CD 格式或 WAV 格式的声音文件。由于其文件尺寸小,音质好,所以在它问世之初还没有什么别的音频格式可以与之匹敌,因而为 MP3 格式的发展提供了良好的条件。直到现在,这种格式还是很受欢迎在被广泛采用,作为主流音频格式的地位难以撼动。

专家点拨: MP3 格式压缩音乐的采样频率有很多种,可以用 64Kb/s 或更低的采样频率节省空间,也可以用 320Kb/s 的标准达到极高的音质。

4. MIDI 格式

MIDI(Musical Instrument Digital Interface)允许数字合成器和其他设备交换数据。MID 文件格式由 MIDI 继承而来。它是一种计算机数字音乐接口生产的数字描述音频文件,扩展名为 mid。

MID 文件并不是一段录制好的声音,而是记录声音的信息,然后告诉声卡如何再现音乐的一组指令。这样一个 MIDI 文件每存 1min 的音乐只用大约 5~10KB。目前,MID 文件主要用于原始乐器作品、流行歌曲的业余表演、游戏音轨以及电子贺卡等。

专家点拨: MID 文件播放的效果完全取决于声卡的档次。MID 格式的最大用处是在计算机作曲领域。MID 文件可以用作曲软件写出,也可以通过声卡的 MIDI 口把外接音序器演奏的乐曲输入计算机里,制成 MID 文件。

5. WMA 格式

WMA(Windows Media Audio)格式是来自于微软的重量级选手,后台强硬,音质要强于 MP3 格式,更远胜于 RA 格式,它和日本 YAMAHA 公司开发的 VQF 格式一样,是以减少数据流量但保持音质的方法来达到比 MP3 压缩率更高的目的,WMA 的压缩率一般都可以达



到 1.18 左右,WMA 的另一个优点是内容提供商可以通过 DRM(Digital Rights Management) 方案如 Windows Media Rights Manager 7 加入防复制保护。这种内置的版权保护技术可以限制播放时间和播放次数甚至于播放的机器等,这对被盗版搅得焦头烂额的音乐公司来说可算是一个福音。

另外,WMA 还支持音频流(stream)技术,适合在网络上在线播放,作为微软抢占网络音乐的开路先锋可以说是技术领先、风头强劲。更方便的是,播放 WMA 格式的音乐不用像 MP3 那样需要安装额外的播放器,而 Windows 操作系统和 Windows Media Player 的无缝捆绑让用户只要安装了 Windows 操作系统就可以直接播放 WMA 音乐,新版本的 Windows Media Player 更是增加了直接把 CD 光盘转换为 WMA 声音格式的功能。在 Windows XP 中,WMA 是默认的编码格式。

WMA 格式在录制时可以对音质进行调节。同一格式,音质好的可与 CD 媲美,压缩率较高的可用于网络广播。目前,WMA 格式在微软的大规模推广下已经得到了越来越多站点的承认和大力支持,在网络音乐领域中直逼 MP3,在网络广播方面也正在瓜分 Real 打下的天下。因此,几乎所有的音频格式都感受到了 WMA 格式的压力。

6. RealAudio 格式

RealAudio 主要适用于在网络上的在线音乐欣赏。Real 文件格式主要有 RA (RealAudio)、RM(RealMedia,RealAudio G2)、RMX(RealAudio Secured)等几种。这些格式的特点是可以随网络带宽的不同而改变声音的质量,在保证大多数人听到流畅声音的前提下,令带宽较富余的听众获得较好的音质。

近来随着网络带宽的普遍改善,Real 公司正推出用于网络广播的、达到 CD 音质的格式。如果用户的 RealPlayer 软件不能处理这种格式,它就会提醒用户下载一个免费的升级包。

3.3.2 音频文件格式的转换

前面介绍了常见的音频文件格式,那么,如果已经有了一张 CD 光盘或是一个 WAV 文件,怎样才能转换成其他的文件格式呢?下面介绍音频文件格式之间的转换问题。

1. CDA 格式转换成 WAV 格式

把 CDA 格式的文件转换成 WAV 格式的声音文件需要使用抓音轨软件,所谓的抓音轨过程就是 CDA 转换成 WAV 格式文件的过程。有些转换软件质量不是很好,在进行抓音轨和声音文件格式转换的时候,声音文件的高频损失很厉害,声音会变得尖锐难听,可以说抓音轨直接影响着最后的声音文件质量。

目前抓音轨软件效果最好的当属 EAC,它可以基本上做到无损抓音轨,这是个免费的自由软件,可以到软件下载网站下载使用。

2. WAV 格式转换成 MP3 格式

在 WAV 格式转换成 MP3 格式的过程中,声音文件的质量关键在于 MP3 的编码器,MP3 编码器之中音质最好的当属 MP3 规范制定者 Fraunhofer IIS 小组开发的 Fraunhofer IIS MPEG Layer3 编码器,这个编码器在 CBR(静态流量编码)上,特别是 128Kb/s 采样频率上的表现是无与伦比的,现在装有 Fraunhofer IIS MPEG Layer3 编码器的音频编辑软件有 Adobe Audition(早期版本软件名称叫 Cool Edit)、Sound Forge 以及功能更加强大的

SAM(Samplitude)等。

3. WAV、MP3 格式转换成 RA 格式

利用 Adobe Audition 软件就可以将 WAV、MP3 格式转换成 RA 格式。操作也十分简便,在 Adobe Audition 软件工作区中选中要转换的文件,利用“另存为”命令即可将文件保存为 RA 格式。

4. WAV、MP3 格式转换成 WMA 格式

要将 WAV、MP3 格式转换成 WMA 格式可以利用 Macromedia 公司功能强大的音频编辑和处理软件 Sound Forge。在 Sound Forge 软件的“另存为”对话框中可以看到 Realmedia 的 RA 格式,另外还有 WMA 选项。操作方法是一样的,也是先在工作区里选择要转换的声音文件,然后另存为 WMA 流媒体格式文件即可。

5. MIDI 格式转换成 WAV、MP3 格式

MIDI 格式的转换有点儿麻烦,大多数的音频编辑软件都不支持 MIDI 格式,这就需要 使用著名的多音轨编辑软件 Samplitude(SAM)。

Samplitude 可以打开并编辑 MIDI 文件,而且在 Samplitude 的音乐输出选项里支持多种多样的音频格式,可以很方便地把 MIDI 文件转换成需要的格式。

专家点拨: Samplitude 的系统资源占用率比较高,请低配置的计算机用户小心使用。

3.4 语音识别技术

语音识别技术是实现人机语音通信所必需的一种关键技术,它的目的是使计算机具有听懂人说话的能力。这也是多媒体技术发展的一个重要方向。

3.4.1 语音识别技术的发展历史

语音识别的研究工作大约开始于 20 世纪 50 年代,当时 AT&T Bell 实验室实现了第一个可识别 10 个英文数字的语音识别系统——Audry 系统。

20 世纪 60 年代,计算机的应用推动了语音识别的发展。这时期的重要成果是提出了动态规划(Dynamic Programming,DP)和线性预测分析技术(Linear Programming,LP),其中后者较好地解决了语音信号产生模型的问题,对语音识别的发展产生了深远影响。

20 世纪 70 年代,语音识别领域取得了突破。在理论上,LP 技术得到进一步发展,动态时间归正技术(Dynamic Time Warping,DTW)基本成熟,特别是提出了矢量量化(Vector Quantization,VQ)和隐马尔可夫模型(Hide Markov Model,HMM)理论。在实践上,实现了基于线性预测倒谱和 DTW 技术的特定人孤立语音识别系统。

20 世纪 80 年代,语音识别研究进一步走向深入,其显著特征是 HMM 和人工神经网络(Artificial Neural Network,ANN)在语音识别中的成功应用。HMM 的广泛应用应归功于 AT&T Bell 实验室 Rabiner 等科学家的努力,他们把原本艰涩的 HMM 纯数学模型工程化,从而为更多研究者了解和认识。ANN 和 HMM 建立的语音识别系统,性能相当。

进入 20 世纪 90 年代,随着多媒体时代的来临,迫切要求语音识别系统从实验室走向实用。许多发达国家如美国、日本、韩国以及 IBM、Apple、AT&T、NTT 等著名公司都为语音识别系统的实用化开发研究投以巨资。



我国语音识别研究工作一直紧跟国际水平,国家也很重视,并把大词汇量语音识别的研究列入 863 计划,由中国科学院声学所、自动化所及北京大学等单位研究开发。我国语音识别技术的研究水平已经基本上与国外同步,在汉语语音识别技术上还有自己的特点与优势,并达到国际先进水平。

3.4.2 语音识别技术的基本原理

首先介绍一下语音识别系统的分类方式及依据。

(1) 根据对说话人说话方式的要求,可以分为孤立字(词)语音识别系统,连接字语音识别系统以及连续语音识别系统。

(2) 根据对说话人的依赖程度可以分为特定人和非特定人语音识别系统。

(3) 根据词汇量多少,可以分为小词汇量、中等词汇量、大词汇量以及无限词汇量语音识别系统。

不同的语音识别系统,虽然具体实现细节有所不同,但所采用的基本技术相似,一个典型语音识别系统的实现过程如图 3-2 所示。

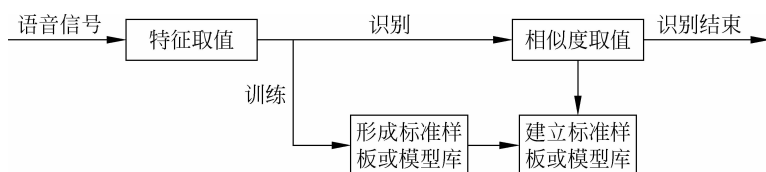


图 3-2 语音识别的基本原理

语音识别分为训练和识别两个阶段。训练阶段是在机器中建立被识别语音的样板或模型库,或者对已经存在的机器中的样板或模型库做特定发音人的适用性修整;在识别阶段,将被识别的语音特征参量提取出来进行模式匹配,相似度最大者即为被识别语音。在大词汇、连续语音识别和口语理解的情况下,使用语言模型对提高识别速度和正确率会起到很大作用。

另外,在语音识别技术中还涉及语音识别单元的选取这一重要环节。选择识别单元是语音识别研究的第一步。语音识别单元有单词(句)、音节和音素三种,具体选择哪一种,由具体的研究任务决定。

单词(句)单元广泛应用于中、小词汇语音识别系统,但不适合大词汇系统,原因在于模型库太庞大,训练模型任务繁重,模型匹配算法复杂,难以满足实时性要求。

音节单元多见于汉语语音识别,主要因为汉语是单音节结构的语言,而英语是多音节,并且汉语虽然有大约一千三百个音节,若不考虑声调,约有 408 个无调音节,数量相对较少。因此,对于中、大词汇量汉语语音识别系统来说,以音节为识别单元基本是可行的。

音素单元以前多见于英语语音识别的研究中,但目前中、大词汇量汉语语音识别系统也在越来越多地被采用。原因在于汉语音节仅由声母(包括零声母 22 个)和韵母(共有 28 个)构成,且声韵母声学特性相差很大。实际应用中常把声母依后续韵母的不同而构成细化声母,这样虽然增加了模型数目,但提高了易混淆音节的区分能力。由于协同发音的影响,音素单元不稳定,所以如何获得稳定的音素单元,还有待研究。

目前,世界各国都加快了语音识别应用系统的研究开发,并已有一些实用的语音识别系

统投入商业运营。在信息处理、教育与商务应用、消费电子应用方面,语音识别技术都已经展现出了它的巨大优势。

就教育领域来讲,语音识别技术最直接的应用就是帮助用户更好地练习语言技巧。以前,用户只是通过简单的模仿来进行语言学习,而无法精确地比较自己发音的差异,应用语音识别技术后,就可以改变这种状况。比如一个美国公司开发的一套 Talk to Me 多媒体教育软件,当用户跟着计算机说完一句话后,计算机同时显示标准发音和用户发音的波形比照图,并给出分数。用户通过比较波形图就可以发现自己在某个发音细节方面的差异,并且可以反复对比倾听来体会这种差异。

专家点拨: 实现人机语音通信,建立一个具有听、说能力的多媒体系统所必需的关键技术除了语音识别技术外,还包括语音合成技术。

3.4.3 语音识别软件介绍

1. ViaVoice

ViaVoice 是 IBM 在 20 世纪 90 年代率先推出的语音识别软件。该系统可用于声控打字和语音导航。只要对着计算机讲话,不用敲键盘即可输入汉字,每分钟可输入 150 个汉字,是键盘输入的两倍,是普通手写输入的 6 倍。该系统识别率可达 95% 以上。并配备了高性能的麦克风,使用便利,特别适合于起草文稿、撰写文章和准备教案,是文职人员、作家和教育工作者的良好助手。

ViaVoice 中文语音识别系统是在 Windows 上使用的中文普通话语音识别听写系统以及相应的开发工具。由于采用连续语音识别技术,汉字输入速度快而且识别率高。可以说, ViaVoice 中文版代表了当前汉语语音识别的最高水平。

2. Dragon NaturallySpeaking

Dragon NaturallySpeaking 是 Scansoft 公司的语音识别软件,它主要应用于英文识别。比其他的语音识别软件更快、更好用,听写、改错和语音控制功能都相当出色。

Dragon NaturallySpeaking 提供的 Vocabulary Optimizer(词汇优化器)可以适应用户的书写风格,而 Performance Assistant(性能助手)则可以根据用户的口述习惯最大程度地提高识别速度。有了这些工具之后,Dragon NaturallySpeaking 能让用户得到很高的识别准确度。

在短短的 5 分钟训练之后,Dragon NaturallySpeaking 的听写准确度就可以达到 90%~95%(根据文档的不同会有所差别)。令人惊异的是,用户可以轻松地用语音来操纵并修改自己的听写结果,而这一点在 IBM 的 ViaVoice 中就不大容易做到。经过一小时的听写、改错和重复训练之后,它的准确度可以上升到 96%~98% 之间。

3.5 基于 Adobe Audition 的数字音频处理技术

Adobe Audition 是著名的音频编辑软件 Cool Edit Pro 的升级版本。本节将介绍使用 Adobe Audition 3.0 进行录制、编辑音频的基本方法。

3.5.1 Adobe Audition 简介

Adobe Audition 是 Adobe 公司发布的一款专业级音频录制、混合、编辑和控制软件。



Adobe Audition 拥有集成的多音轨和编辑视图、实时特效、环绕支持、分析工具、恢复特性和视频支持等功能,为多媒体开发人员提供全面集成的音频编辑和混音解决方案。它支持工业标准音频文件格式,包括 WAV、AIFF、MP3、MP3PRO 和 WMA 等,还能够利用高达 32 位的位深度来处理文件,取样速度超过 192kHz,从而能够以最高品质的声音输出磁带、CD、DVD 或 DVD 音频。

在 Adobe Audition 3.0 版本中又新增了不少的功能,如支持 VSTi 虚拟乐器、增强的频谱编辑器、增强的多轨编辑、iZotope 授权的 Radius 时间伸缩工具、新增吉他系列效果器、可快速缩放波形头部和尾部、增强的降噪工具和声相修复工具、更强的性能、波形编辑工具等。

启动 Adobe Audition 3.0 中文版,界面默认为“多轨界面”模式(单击“窗口”→“操作界面”命令可更改界面模式),如图 3-3 所示。

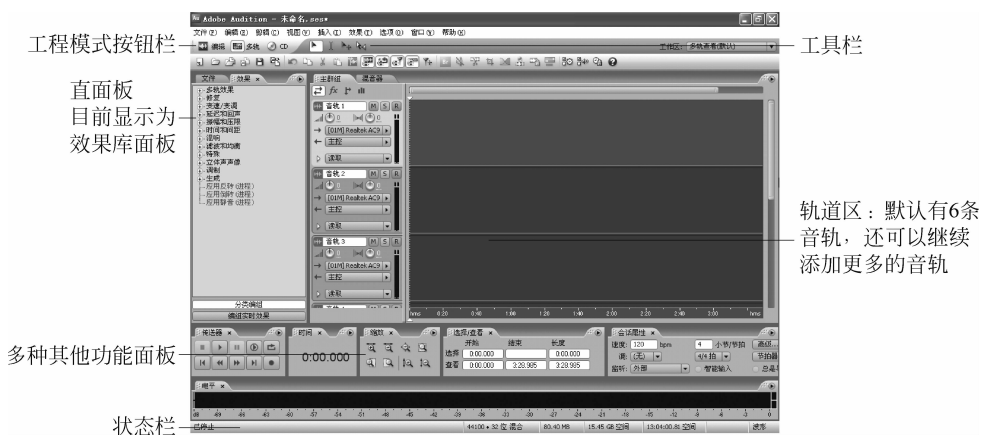





图 3-3 Adobe Audition 3.0 中文版启动界面

1. 工程模式按钮栏

工程模式按钮栏位于菜单栏下方。单击“编辑”按钮可以进入单轨编辑查看模式,进行音频文件的破坏性编辑处理;单击“多轨”按钮可以进入多轨查看模式,可进行录音和多轨混音工作;单击“光盘”按钮可进入光盘编排模式,适合 CD 刻录前的曲目安排。单击最右侧的“操作界面”下拉列表还可以选择更多的工程模式。

2. 工具栏

工具栏位于工程模式按钮栏右侧,在不同的模式下,工具栏中会显示相对应的工具按钮。

3. 主面板

主面板分为主群组“轨道区”和左侧的“直面板”两部分。主群组“轨道区”默认情况下会显示 6 条音轨,每一条音轨都可以独立设置。选择“插入”→“添加音频”命令可以添加更多的音轨。

在左侧的“直面板”中选择“文件”选项卡,可打开素材库面板,查看所有被导入 Audition 的文件;选择“效果”选项卡,可打开效果库面板,查看调用 Audition 效果器或 Audition 支持的第三方效果器。

4. 多种其他功能面板

默认显示“传送控制器”面板、“时间”面板、“缩放控制”面板、“选择/查看”面板、“会话属性”面板及“电平表”面板。在不同的工程模式下显示的功能面板也不同。

5. 状态栏

显示“当前音轨状态”、“采样格式”、“文件大小”、“剩余空间”、“剩余时间”、“会话属性”等信息。

3.5.2 使用 Adobe Audition 前的相关设置

为了更好地使用 Adobe Audition 3.0 进行录音、混音及其他操作,首先要对 Adobe Audition 进行一些必要设置。

1. 硬件音频设备设置

ASIO 声卡可以摆脱 Windows 操作系统对硬件的集中控制,实现在音频处理软件与硬件之间进行多通道传输的同时,将系统对音频流的响应时间降至最短,提高声卡音质。在 Adobe Audition 中使用 ASIO 声卡驱动还可以实现录音效果监听,即以前只有专业声卡才能实现的“听湿录干”的效果——唱的时候在耳机里监听到带效果的声音,但录到计算机里的录音文件是不带效果的,方便以后的后期效果处理。如果计算机使用的是普通 AC97(集成声卡),可以到 <http://www.asio4all.com> 网站下载一个 ASIO 驱动。并非所有声卡都能适用非声卡制造商提供的 ASIO 驱动。当然,使用 Adobe Audition 3.0 进行录音也不是非要使用 ASIO 声卡,使用计算机中的原有声卡同样可以进行录音工作。

下面简单介绍一下该设备的设置方法。

(1) 选择“编辑”→“设置音频设备”命令,打开“设置音频设备”对话框。

(2) 在对话框中,分别在“编辑界面”、“多轨界面”、“环绕声编码器”选项卡中选择音频驱动为 ASIO4ALL v2,如图 3-4 所示。



图 3-4 设置音频设备



2. 设置录音设备


- (1) 选择“选项”→“Windows 录音控制”命令,打开“录音控制”对话框。
- (2) 在“录音控制”对话框中选择“麦克风”,并调整其音量设置。单击“高级”按钮,在“麦克风的高级控制”对话框中选中 1 Mic Boost 复选框,以增强麦克风音量,如图 3-5 所示。



图 3-5 “麦克风的高级控制”对话框

3. 启用 DirectX 效果

在安装 Adobe Audition 3.0 软件支持的第三方效果插件后,必须先启用 DirectX 效果才能够正常使用。

- (1) 单击工程模式按钮栏中的“编辑”按钮,切换到“编辑查看”模式。
- (2) 选择“效果”→“启用 DirectX 效果”命令。
- (3) 在弹出的警告对话框中单击“确定”按钮。开始刷新效果列表,如图 3-6 所示。

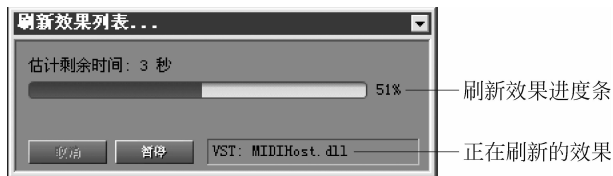


图 3-6 “刷新效果列表...”对话框

- (4) 效果刷新完成后,安装插件的效果可以在“效果”→DirectX 里面找到。
- (5) 以后再安装的插件,必须执行“效果”→“刷新效果列表”命令才能在效果列表中找到。

3.5.3 使用 Adobe Audition 录音

录音是音频所有后期制作加工的基础,在录音过程中出现的问题靠后期加工是无法完美地弥补的,有时甚至是无能为力的。因此,在录音时就要注重每一个细节,如果原始的录音有较大问题,还是重录为好。

在单轨模式下的录音方法相对简单,只需单击“传送带”面板中的“录音”按钮,在弹出的

“新建波形”对话框中设定声音的采样率、声道和分辨率(即采样精度或称为量化级)就可以开始录音了。

在本节中以在多轨模式下根据伴奏录制自唱的歌曲为例介绍录音的整个过程。

专家点拨: 录音时要关闭音箱,通过耳机来听伴奏。录制前,一定要调节好总音量及麦克风音量,如果麦克风音量过大,会导致录出的波形成了方波,这种波形的声音是失真的,无论编辑水平多么高超,也不可能处理出令人满意的结果。

(1) 开启 Adobe Audition 3.0 多轨模式,在默认情况下,软件自动创建了一个采样率为 44 100 的会话工程,如果要创建其他采样率的会话工程,可以选择“文件”→“新建会话”命令,在弹出的“新建会话”对话框中选择合适的采样率,如图 3-7 所示。



图 3-7 “新建会话”对话框

专家点拨: 如果录制的是歌曲,以 44 100 采样率为好,如果录制的只是一般的语音,可选择相对低一些的采样率。

(2) 在“音轨 1”的音轨区右击,选择“插入”→“音频”命令,在弹出的“插入音频”对话框中选择伴奏音乐,单击对话框右下角的“播放”按钮还可以试听音乐,如图 3-8 所示。

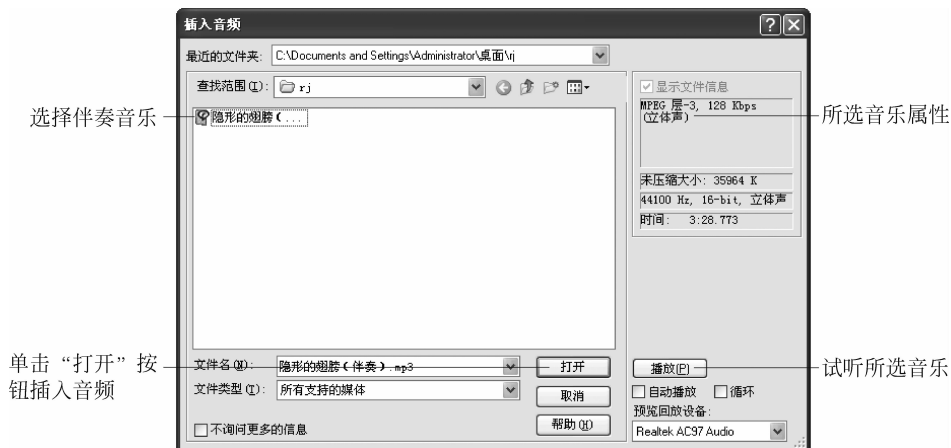
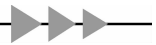


图 3-8 “插入音频”对话框

专家点拨: 如果以后需要将此伴奏音乐和自录的歌曲进行混音,要留意此时导入的伴奏音乐是否和创建的新工程采样率一致,以免在后期的混音过程中出现麻烦。

(3) 单击“打开”按钮,在几秒钟的读取音频数据过程后,文件被插入到“音轨 1”中,单击“音轨 1”标题框,更改音轨名称为“伴奏音”。



(4) 更改“音轨 2”名称为“录音”，单击此音轨中的“输入”项，在弹出的选项中选择“立体声”→[01S]Realtek AC97 Audio 1 端口，如图 3-9 所示。

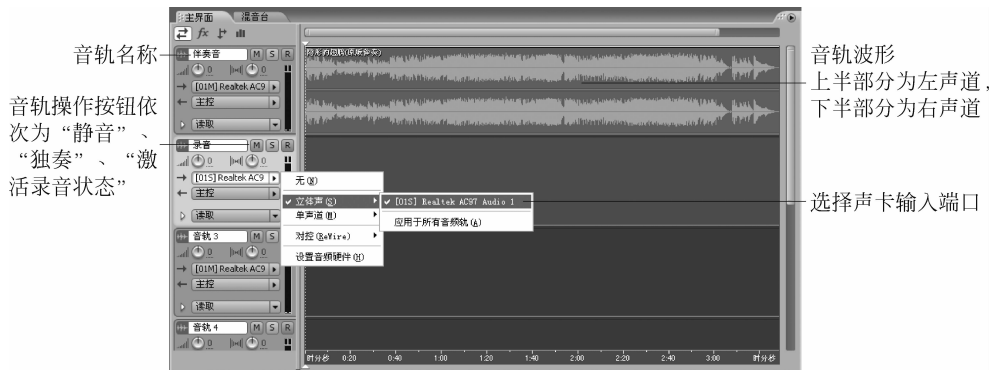


图 3-9 设置“录音”音轨输入端口


(5) 在“录音”音轨上单击“激活录音状态”按钮 **R**，在弹出的“保存会话为”对话框中设定保存的路径和名称，单击“确定”按钮，如图 3-10 所示。




图 3-10 “保存会话为”对话框


接下来将设置启用“听湿录干”效果，如果不想用此效果，可以直接跳至第(8)步。

(6) 选择“选项”→“监听”→“内部混合”→“智能化监视”命令，如图 3-11 所示。此时就已经能从耳机中听到自己的声音了。

(7) 选择“窗口”→“效果格架”命令，在弹出的对话框中选择相应的效果，如添加“修复”→“适应性降噪”效果，调整各项指标至最佳效果，单击降噪指标项上方的“保存”按钮  保存设置，如图 3-12 所示。关闭对话框。

(8) 保持激活“录音”轨道，将呈黄色虚线的播放指针移至音轨最左侧，单击底部“传送器”面板中的“录音”按钮 ，开始跟着伴奏边唱边录。

专家点拨：尽量要保持室内环境的安静，以免出现其他噪声。

(9) 录制结束后可单击“停止”按钮  或按空格键停止录音。声音文件被自动保存成 WAV 格式的文件。单击“传送器”面板中的“播放”按钮进行试听，如图 3-13 所示。

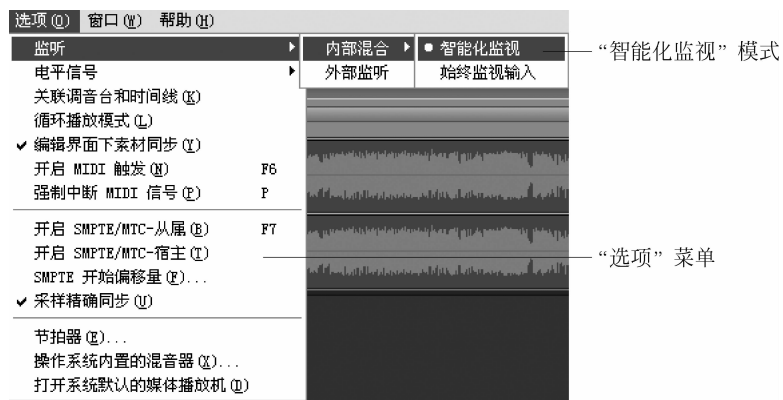


图 3-11 设置“智能化监视”

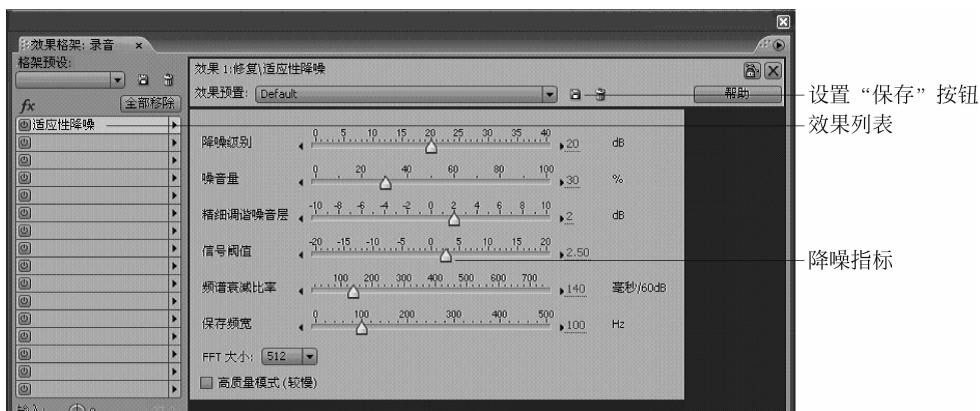


图 3-12 “效果格架：录音”对话框

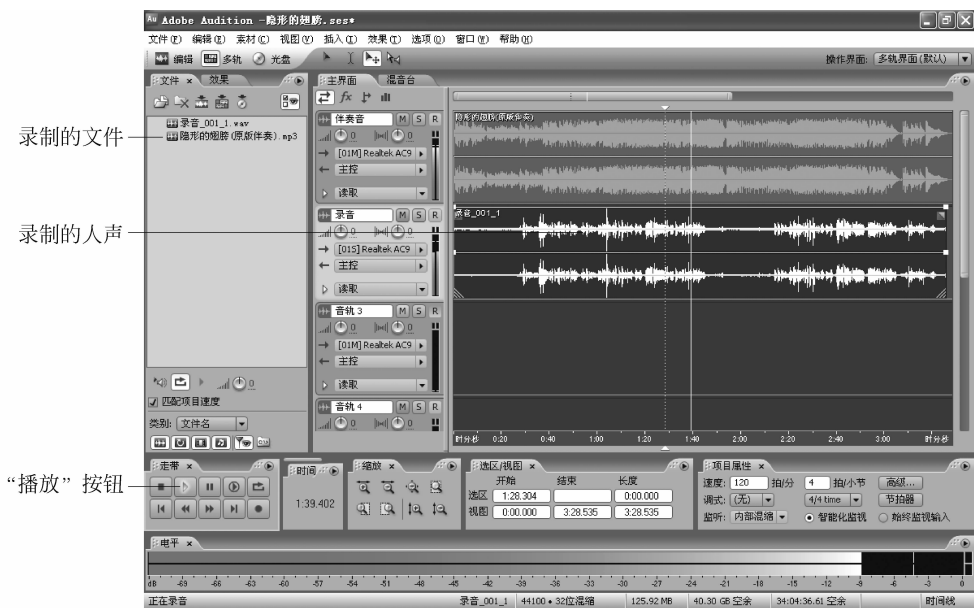
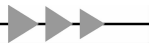


图 3-13 试听录制的声音



(10) 最后保存会话即可。

3.5.4 使用 Adobe Audition 编辑音频



在 Adobe Audition 3.0 中有两种编辑处理音频的方式,一种是在单轨编辑界面下进行编辑处理,这种处理一般是破坏性编辑,即直接对音频源文件进行编辑。另一种方式是在多轨界面中编辑处理,这种编辑只体现在多轨模式下的变化,并不改变音频源文件内容。

下面介绍使用 Adobe Audition 在单轨编辑模式下对声音的基本编辑处理方法,如剪切、复制、粘贴、混合音频等。

1. 准确选择需要编辑的波形

音频文件的编辑工作相比其他文件更加精细,特别是有节奏的音频,差半拍就会让人觉得不舒服,因此在编辑声音文件时更需要精确地选择音频的波形。

(1) 在“编辑查看”模式下打开要编辑处理的音频。

(2) 单击“传送器”面板中的“播放”按钮或按空格键开始监听音频的播放,当播放到关键的波形时可单击“快捷栏”中的“添加标记”按钮 ,或按 F8 键在波形上添加一个标记。如果波形选区要求非常精确,可以先单击“缩放控制”面板中的“水平放大”按钮 ,放大波形视图再反复调整。

专家点拨: 默认界面中是不显示“快捷栏”的,选择“视图”→“快捷栏”→“显示”命令,可将常用工具的快捷栏显示出来。

(3) 选择波形除了可以用鼠标直接拖选两个标记间的波形,也可以先记下每个标记的时间,然后将选区的开始和结束时间输入到“选择/查看”面板的相应文本框中,如图 3-14 所示。

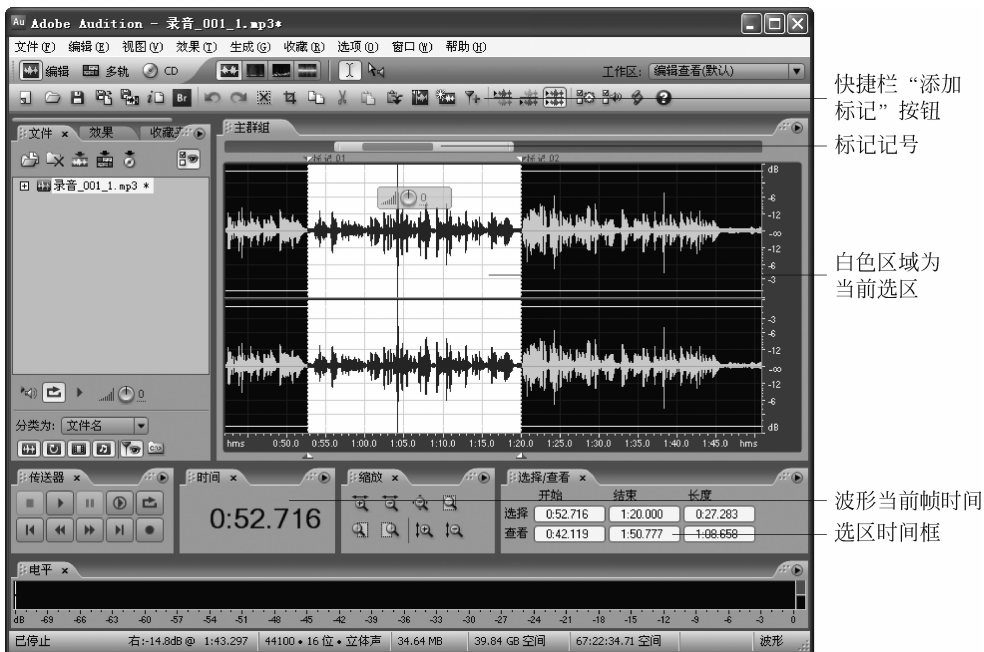


图 3-14 选择需要编辑的波形

(4) 如果要选中右声道中的波形,则可以将光标移到波形窗口下方边界时,光标显示 R 的时候拖动鼠标。同样地,如果要选中左声道中的波形,则可以将光标移到波形窗口上方边界时,光标显示 L 的时候拖动鼠标。

2. 编辑波形


选择好波形选区后,可对选区进行剪切、删除、复制、粘贴等操作。

(1) 选择“编辑”→“剪贴板设置”菜单中的一个空剪贴板,如图 3-15 所示。



图 3-15 选择一个剪贴板

专家点拨: Adobe Audition 3.0 共提供了 5 个剪贴板,即用户在剪贴板中最多可以同时保存 5 段波形,其中最后一项 Windows 是系统剪贴板,在这个剪贴板中保存的波形可以在其他音频编辑软件中使用。

(2) 单击“快捷栏”中的“剪切”按钮,剪切被选中的一段波形。

(3) 将播放指针移到合适位置,单击“快捷栏”中的“粘贴”按钮就可以完成波形的移动操作。如果要粘贴的内容不在当前选择的剪贴板中,必须先在“剪贴板设置”中更改剪贴板的位置,再进行粘贴。

选择“编辑”→“混合粘贴”命令,还可以将剪贴板中的波形与当前帧处的波形进行混合,如图 3-16 所示。



图 3-16 “混合粘贴”对话框

3.5.5 使用 Adobe Audition 优化声音效果

Adobe Audition 3.0 在声音效果的优化方面功能强大,除了可以使用软件自带的很多



专业的声音效果编辑优化之外,还可以安装软件所支持的第三方效果进行优化处理。下面以录制的“隐形的翅膀”为例简单介绍 Adobe Audition 3.0 部分声音效果的使用。

1. 音量调整

- (1) 打开“隐形的翅膀”会话工程,双击“录音”音轨波形进入单轨编辑界面。
- (2) 选择需要调整的波形,选择“效果”→“振幅与压限”→“振幅/淡化(进程)”命令。
- (3) 在弹出的对话框的“预设”列表框中选择要处理的分贝数,也可以直接拖动左侧“常量”选项卡中声道的滑块或在分贝框中填入分贝数字,如图 3-17 所示。



图 3-17 “振幅/淡化”对话框“常量”选项卡

- (4) 单击“试听”按钮进行试听,如不满意可重新调整,完成后单击“确定”按钮开始处理。

2. 淡入淡出效果

- (1) 选择需要调整的波形,选择“效果”→“振幅与压限”→“振幅/淡化(进程)”命令。
- (2) 在弹出的“振幅/淡化”对话框中选择淡入淡出的“预设”选项,也可以选择对话框中的“渐变”选项卡,自由设置淡入淡出参数,如图 3-18 所示。

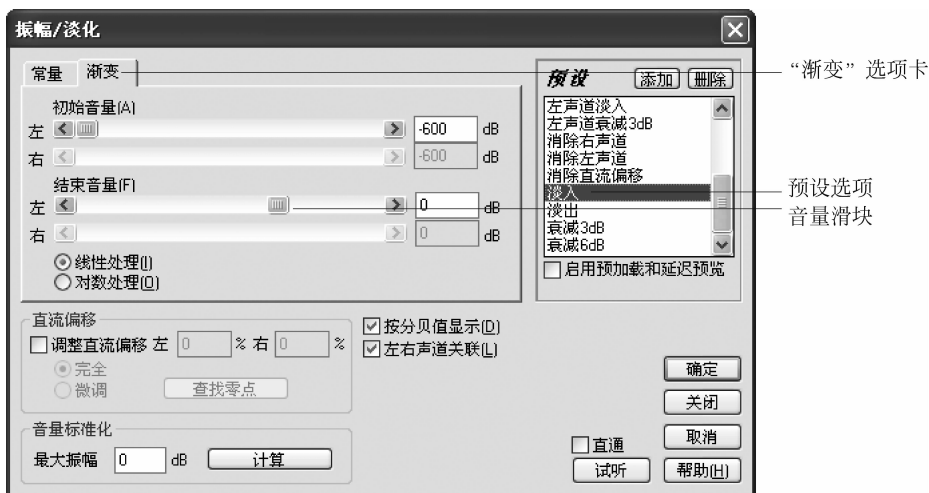


图 3-18 “振幅/淡化”对话框“渐变”选项卡

3. “降噪”处理

(1) 选取一段没有人声的波形, 右击, 在弹出的快捷菜单中选择“采集降噪预置噪声”命令, 如图 3-19 所示。

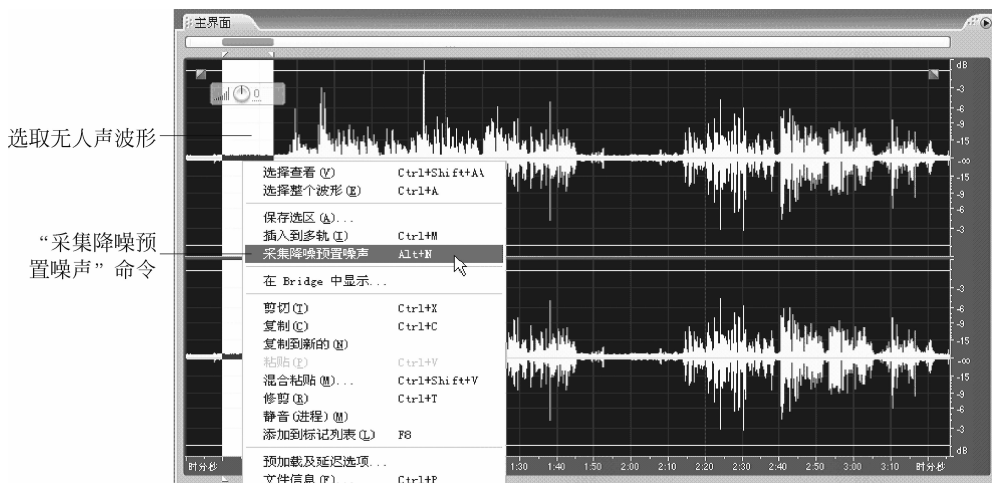


图 3-19 获取噪声特征

(2) 在关闭“获取噪声特性”提示对话框及几秒钟的数据读取后, 完成噪声的采样工程。

(3) 单击左侧直面板中的“效果”标签, 在树状菜单中双击“修复”→“降噪器(进程)”命令, 弹出“降噪器”对话框, 如图 3-20 所示。

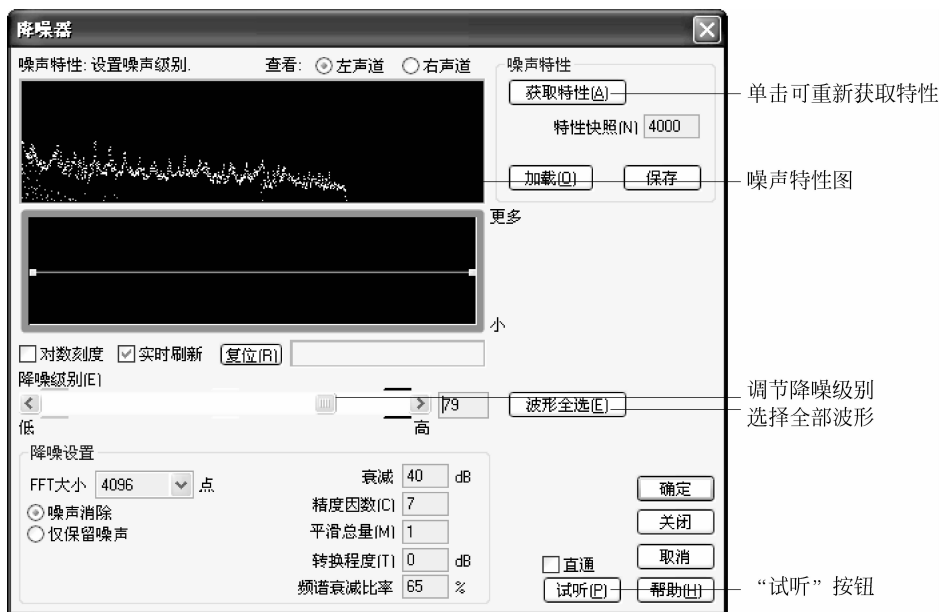


图 3-20 “降噪器”对话框

(4) 单击“波形全选”按钮选择全部波形, 调节降噪级别, 注意不要调得太高, 以免声音失真, 单击“试听”按钮试听后再调整到最佳状态, 单击“确定”按钮开始降噪, 如图 3-21 所示。

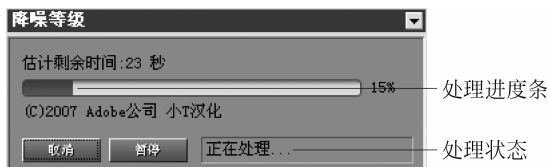


图 3-21 降噪处理

如果降噪完成后感觉降噪太过,可以撤销操作重新进行降噪处理。

4. 变速效果

(1) 在效果选项卡的树状菜单中双击“时间和间距”→“变速(进程)”命令。打开“变速”对话框。

(2) 在“转换”下拉列表中选择合适的升降调,“#”号为升调,“b”号为降调。前面的数字代表升降的度数,以半度为一个单位,如 1# 为升半度,2# 为升一度,以此类推,如图 3-22 所示。




图 3-22 “变速”对话框

(3) 设置“精度”、“变速模式”和“变调变速设置”等选项。单击“试听”按钮进行试听,满意后单击“确定”按钮进行升降调处理。

5. 添加混音效果

除了在音轨编辑模式下添加效果之外,在多轨模式下同样可以添加声音效果,唯一的区别是多轨模式下添加的效果不会影响到该音轨的声音源文件。具体操作如下所述。

(1) 单击“多轨”界面按钮,返回到“多轨”模式下。

(2) 单击“主群组”顶部的“效果”按钮 ,在“录音”音轨中单击效果框右侧的小三角按钮,在弹出的菜单列表中选择“混响”→“房间混响”命令,如图 3-23 所示。

(3) 在弹出的“效果格架: 录音”对话框中设置“房间混响”的具体参数,如图 3-24 所示。

(4) 效果添加完以后就可以将人声和伴奏混缩在一起了。选中两条音轨,选择“编辑”→“混缩为新文件”→“所选范围的主输出—立体声”命令,开始进行“混缩”处理。

(5) “混缩”完成后自动跳转到混缩文件的单轨编辑界面,如果效果满意可选择“文件”→“另存为”命令将声音保存为其他格式的音频文件,如 mp3PRO 格式,单击下方的“选项”按钮还可以设置具体的编码选项等内容,如图 3-25 所示。单击“保存”按钮完成。

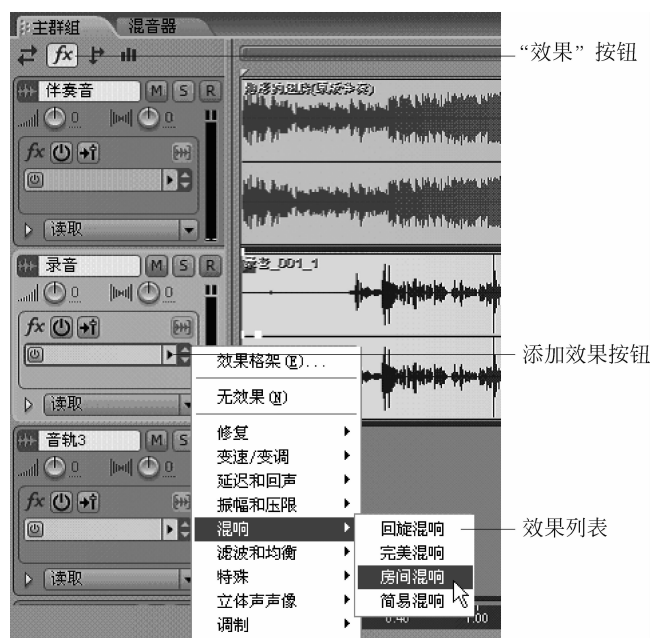


图 3-23 添加“房间混响”效果

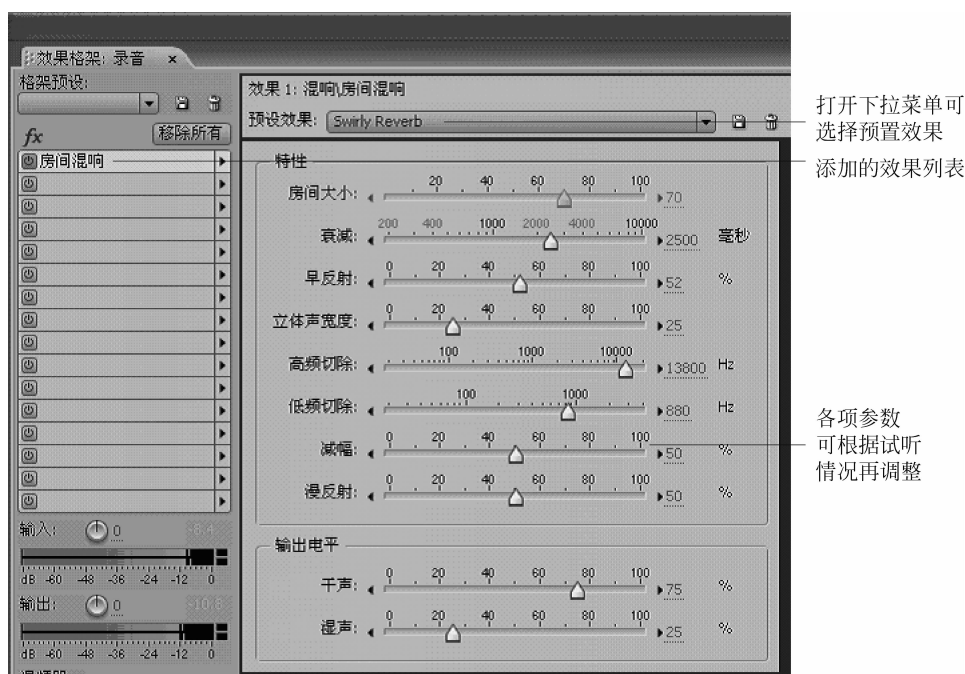


图 3-24 设置“房间混响”效果参数

专家点拨：这里添加的混响效果只在多轨模式下起作用,当两个音轨混缩成一个声音文件时,混响效果也会被混缩进最后的混缩文件中,但在单独的人声源文件中却不会有任何变化。

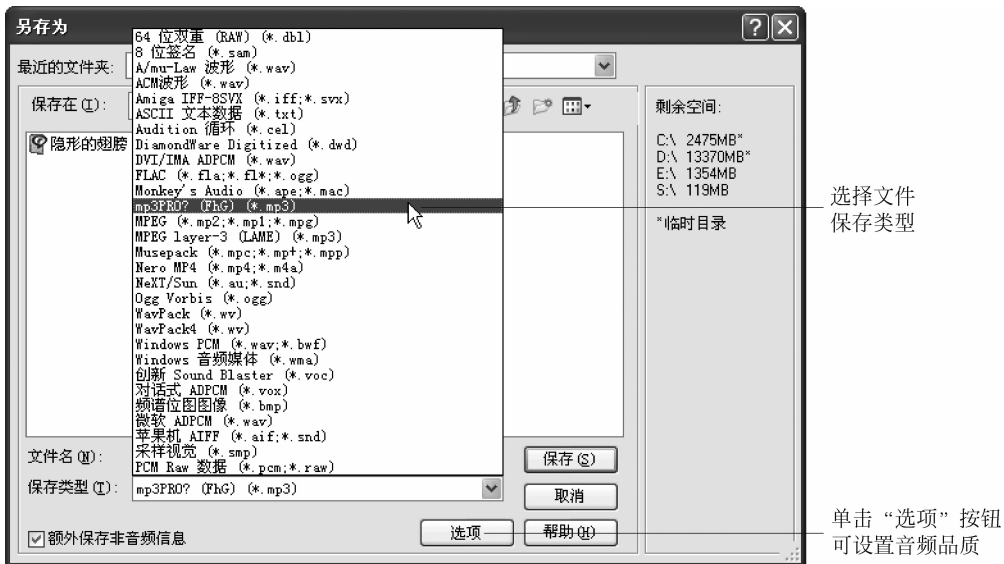
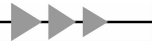


图 3-25 录制音频“另存为”对话框

习 题

1. 选择题

(1) ()是每秒振动的次数,它的单位为赫兹(Hz),是以一个名叫海里奇 R. 赫兹的人命名的。

- A. 振幅 B. 周期 C. 频率 D. 音调

(2) ()是对模拟音频信号的幅度进行数字化,它决定了模拟信号数字化以后的动态范围。它越高,信号的动态范围越大,数字化后的音频信号就越可能接近原始信号,但所需要的存储空间也越大。

- A. 采样频率 B. 量化位 C. 声道数 D. 编码算法

(3) 它允许数字合成器和其他设备交换数据,它是一种计算机数字音乐接口生产的数字描述音频文件。这种音频格式为()。

- A. WAV B. MP3 C. WMA D. MIDI

(4) ViaVoice 是一个()类型的软件。

- A. 音频格式转换 B. 语音识别 C. 音频编辑 D. 英语翻译

2. 填空题

(1) 从听觉角度看,声音具有_____、_____和_____三要素。声音质量的高低主要取决于这三要素。

(2) 把模拟信号转换成数字信号的过程称为模/数转换,它主要包括_____、_____和_____三个步骤。

(3) 如果以 CD 质量(44.1kHz 的采样频率、16 位的量化位数、双声道立体声)记录一首时长 5min(300s)的音乐,那么数据量为_____。

(4) 一个 CD 音频文件是一个 cda 文件,但这只是一个_____,并不是真正包含的声音信息,因此不论 CD 音乐的长短,在计算机上看到的 cda 文件都是_____字节长。

上机练习

练习 1 使用 Adobe Audition 处理声音素材

在 Internet 或者素材光盘上获取一段声音素材,使用 Adobe Audition 对其进行编辑和处理。

制作要点提示:

- (1) 在 Adobe Audition 中打开声音素材。
- (2) 裁剪声音素材。
- (3) 对声音进行淡入、淡出效果的处理。
- (4) 将声音保存为其他格式并进行比较。

练习 2 使用 GoldWave 软件进行录音

GoldWave 也是一个功能强大并且体积小巧的数字音频编辑软件,请使用它进行录音练习。

制作要点提示:

(1) 启动 GoldWave 软件,新建一个声音文件,设置声音的“声道”、“采样率”、“持续时间”等选项,如图 3-26 所示。



图 3-26 “新建声音”对话框

(2) 双击任务栏右下角的“声音”图标,设置好录音选项及麦克风的音量。

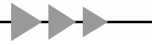
(3) 单击“控制器”窗口中的“录音”按钮  开始录音,完成后单击“停止”按钮  结束录音。

练习 3 使用 GoldWave 制作卡拉 OK 伴奏音乐

利用 GoldWave 软件,将带有原唱的歌曲文件中的人声消除掉,得到卡拉 OK 伴奏音乐。

制作要点提示:

- (1) 启动 GoldWave 软件,打开带有原唱的歌曲文件。
- (2) 选择“效果”→“立体声”→“减少人声”命令,在弹出的对话框中设置,通过试听反复



调整,如图 3-27 所示。



图 3-27 “减少人声”效果对话框

(3) 单击“确定”按钮应用效果。

练习 4 使用“全能音频转换通”软件进行音频格式转换

全能音频转换通是一款音频文件格式转换软件。典型的应用如 WAV 转 MP3,MP3 转 WMA,WAV 转 WMA,RM(RMVB)转 MP3,AVI 转 MP3,RM(RMVB)转 WMA 等。

在 Internet 或者素材光盘上获取一些 MP3 音乐,练习使用“全能音频转换通”软件进行音频格式的批量转换。

制作要点提示:

(1) 启动全能音频转换通 V1.2 软件,如图 3-28 所示。单击“添加文件”按钮。

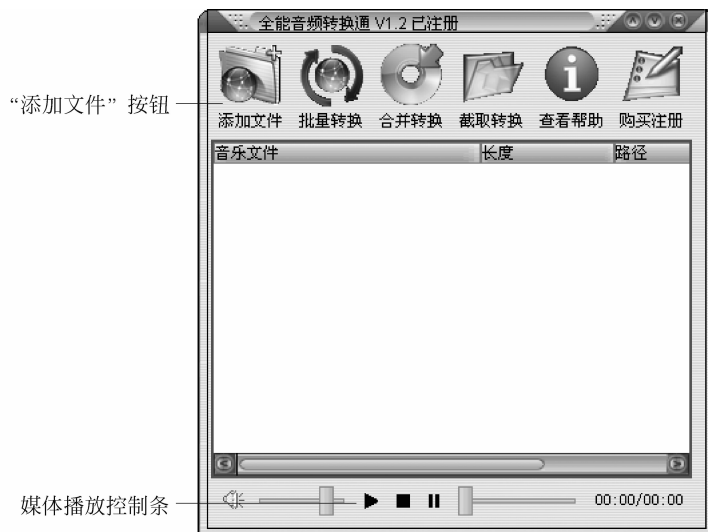


图 3-28 全能音频转换通 V1.2 界面

(2) 在弹出的“打开”对话框中选择多个需要转换的音频文件(文件格式不需要完全一样)。

(3) 选中其中的一首歌曲,单击播放控制条上的“播放”按钮进行试听。可以选择不需要转换的曲目,按 Delete 键或右击选择“删除所选”命令进行删除,如图 3-29 所示。



图 3-29 删除列表中多余曲目

(4) 单击“批量转换”按钮,进入“批量转换文件格式”窗口,对“输出格式”、“选择编码器”、“输出质量”等参数进行设置,并设定“对于重名文件”的操作方法及转换后的目标文件夹,如图 3-30 所示。最后单击“开始转换”按钮。



图 3-30 “批量转换文件格式”窗口

(5) 转换的速度根据文件的多少及转换的质量会有所不同,一般转换一首歌曲需要的时间不会超过歌曲长度的一半。转换完成后,单击右上角的“关闭”按钮即可回到软件主界面。